



(12) **EUROPÄISCHE PATENTANMELDUNG**

(43) Veröffentlichungstag:
05.09.2001 Patentblatt 2001/36

(51) Int Cl.7: **G10L 21/02**

(21) Anmeldenummer: **01102129.2**

(22) Anmeldetag: **01.02.2001**

(84) Benannte Vertragsstaaten:
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU
MC NL PT SE TR**
Benannte Erstreckungsstaaten:
AL LT LV MK RO SI

(71) Anmelder: **Volkswagen Aktiengesellschaft
38436 Wolfsburg (DE)**

(72) Erfinder:
• **Schultz, Jürgen
38554 Weyhausen (DE)**
• **Schaaf, Klaus, Dr.
38116 Braunschweig (DE)**

(30) Priorität: **02.03.2000 DE 10010037**

(54) **Verfahren zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen**

(57) Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen, bei dem das technische Problem, für eine Wiedergabe eines gestörten Sprachsignals eine möglichst naturgetreue Wiedergabe zu ermöglichen, durch ein Verfahren gelöst wird, bei dem oberhalb einer Grenzfrequenz ω_0 mindestens zwei benachbart angeordnete Frequenzanteile ω_{fa1} , ω_{fa2} , ... mit erhöhter Amplitude im Sprachsignal be-

stimmt werden und bei dem die Grundfrequenz ω_g des Sprachsignals als Frequenzdifferenz zwischen den mindestens zwei benachbarten Frequenzanteilen ω_{fa1} , ω_{fa2} , ... bestimmt wird und bei dem mit Hilfe der ermittelten Grundfrequenz ω_g und des Sprachsignals der tieffrequenten Frequenzbereich unterhalb der Grenzfrequenz ω_0 rekonstruiert wird.

Die Erfindung betrifft auch eine Vorrichtung zur Durchführung des vorgenannten Verfahrens.

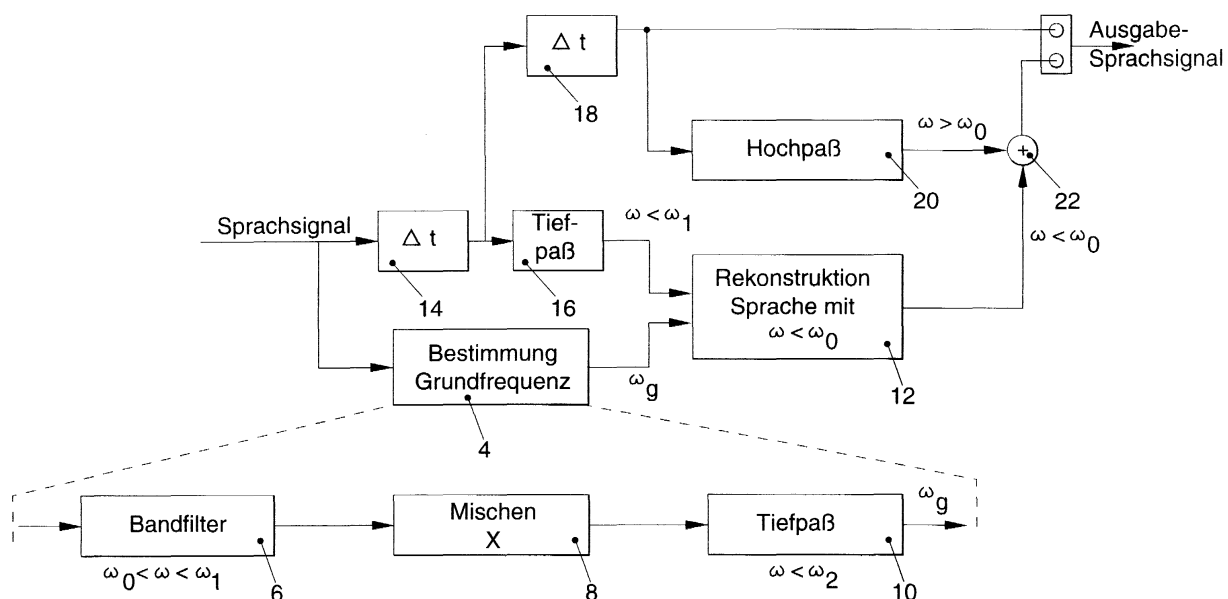


FIG. 7

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren und eine Vorrichtung zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen.

[0002] Im Stand der Technik der digitalen Verarbeitung von Sprachsignalen mit einem hohen Lärmpegel im tieffrequenten Bereich wird das Signal dadurch verbessert, daß entweder Störanteile herausgefiltert werden oder sehr stark gestörte Frequenzbereich aus dem Signal vollständig herausgefiltert werden.

[0003] Aus der US 5,842,160 A ist ein Verfahren zur Verbesserung der Qualität einer digitalen Sprachübertragung bekannt, bei dem verschiedenen Frequenzbändern je nach Energiegehalt verschiedene Datenmengen zugeordnet werden. Durch die Art der Kodierung und Übertragung entstehen niederenergetische Signalebereiche, die zu Lücken im empfangenen Signalspektrum führen. Diese Lücken werden durch synthetisch aus den vorhandenen Daten gewonnenen Signale gefüllt, so daß ein natürlicher klingendes Sprachsignal erreicht wird.

[0004] Aus der US 4,091,237 A ist ein Verfahren zur Ermittlung der Stimmgrundfrequenz eines digitalen Sprachsignals in Echtzeit bekannt. Speziell für Signale mit einem eingeschränkten Frequenzbereich, wie Telefonsignale, und mit einem hohen Störgeräuschanteil wird das Sprachsignale verbessert, indem Störgeräusche ausgefiltert werden. Das Signal wird durch eine Mehrzahl von Bandpaßfiltern aufgesplittet und ein entsprechendes Histogramm gebildet, aus dem die Stimmgrundfrequenz extrahiert wird. Ist die Grundfrequenz bekannt, können Störgeräusche daran erkannt werden, daß sie in keinem harmonischen Verhältnis zur Grundfrequenz stehen. Das zuvor beschriebene Verfahren dient dazu, die für eine Stimme charakteristische Grundfrequenz zu bestimmen.

[0005] Weiterhin ist aus der DE 37 33 983 ein Verfahren zum Dämpfen von Störsignalen in einem Hörgerät bekannt, bei dem das Signal digitalisiert und in einzelne Frequenzbereiche aufgeteilt wird. Frequenzbereiche mit bestimmten Charakteristika, wie schnelle oder sehr langsame Spektralverteilungsänderungen, werden gedämpft und/oder es werden die Grenzfrequenzen verschoben. Das so gereinigte Signal wird in synthetische Sprachsignale umgewandelt.

[0006] Den zuvor beschriebenen Verfahren und den damit verbundenen Vorrichtungen liegt der Nachteil zugrunde, daß das Sprachsignal gar nicht oder nur in unzureichender Form rekonstruiert wird, um ein möglichst natürliches Ausgangssprachsignal zu erzeugen.

[0007] Die zuvor dargestellten Verfahren können unter anderem bei der digitalen Sprachverstärkung (digital voice enhancement - DVE) eingesetzt werden. Beispielsweise sind oberhalb jeder Sitzreihe in einem Kraftfahrzeug zwei Mikrophone angebracht, so daß es bspw. allen Fahrzeuginsassen ermöglicht wird, sich an einem Telefongespräch zu beteiligen. Das System überträgt

dazu die Sprache, die vorn durch das Mikrophon aufgenommen wurde, auf die hinteren Serienlautsprecher und umgekehrt. Das System ist somit voll mit dem Freisprechtelefon und dem Radio/CD/Navigationsgerät gekoppelt. Es verbessert insbesondere bei schneller Fahrt die Verständigung innerhalb des Fahrzeuges deutlich.

[0008] Der Pegel des Fahrzeuginnengeräusches steigt zu tiefen Frequenzen sehr stark an, so daß die Sprache dort vom Lärm überdeckt wird. Um durch das DVE-System möglichst wenig Umgebungslärm zu übertragen, denn dadurch würde der Innenlärmpegel unnötig erhöht, werden bei einem Teil der oben beschriebenen Verfahren alle Frequenzen je nach Geschwindigkeit unterhalb von bspw. 200 bis 500 Hz abgeschnitten. Die Folge ist, daß die Sprachgrundfrequenz und die ersten Vielfachen (Harmonischen) im übertragenen Signal fehlen. Die Sprache klingt somit telefonartig, da typischer Weise ein Telefonnetz eine Klangübertragung nur oberhalb von 350 Hz ermöglicht.

[0009] Neben der Nutzung eines Freisprechtelephons kann mit den Verfahren auch die Sprachverständigung innerhalb des Fahrzeuges durchgeführt werden. Dabei ist jedoch eine optimale Klangqualität erforderlich, um eine Akzeptanz bei den Käufern zu erzielen.

[0010] Insbesondere bei den Verfahren, die die Sprache von Störgeräuschen befreien, z. B. spektrale Subtraktion oder Kohärenzfiltern, kommt es dazu, daß die Varianz der Frequenzkomponente von Rauschen in die Größenordnung der Leistung des Sprachsignals kommt. Somit ist eine effektive Rauschunterdrückung nicht mehr möglich und die angewendeten Verfahren greifen nicht mehr.

[0011] Der Erfindung liegt daher das technische Problem zugrunde, das aus dem Stand der Technik bekannte Verfahren sowie die zugehörige Vorrichtung zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen dahingehend weiterzubilden und auszugestalten, daß für eine Wiedergabe des gestörten Sprachsignals eine möglichst naturgetreue Wiedergabe ermöglicht wird.

[0012] Das zuvor aufgezeigte technische Problem wird durch ein Verfahren mit den Merkmalen des Anspruches 1 gelöst. Zunächst werden oberhalb einer Grenzfrequenz mindestens zwei benachbart angeordnete Frequenzanteile mit erhöhter Amplitude im Sprachsignal bestimmt. Danach wird die Grundfrequenz des Sprachsignals als Frequenzdifferenz zwischen den mindestens zwei benachbarten Frequenzanteilen bestimmt. Schließlich wird mit Hilfe der ermittelten Grundfrequenz und des Sprachsignals der tieffrequente Frequenzbereich unterhalb der Grenzfrequenz rekonstruiert. Das somit erzeugte synthetische Sprachsignal kann dann über eine Wiedergabevorrichtung direkt wieder ausgegeben werden oder für ein späteres Aussenden gespeichert werden.

[0013] Mit anderen Worten werden tieffrequente Signalanteile des Sprachsignals synthetisch erzeugt, also rekonstruiert, und den restlichen aufgenommenen

Sprachsignal zugemischt. Die Rekonstruktion der tiefrequenten Sprachanteile geschieht dabei auf der Grundlage der nicht ausgefilterten Sprachsignale. Dazu wird ausgenutzt, daß die tieffrequenten Sprachanteile von höherfrequenten Anteilen der Harmonischen begleitet sind, so daß sich die fehlenden Anteile aus dem verbleibenden Signal abschätzen lassen.

[0014] In bevorzugter Weise werden neben der Grundfrequenz auch die Frequenzen der unterhalb der Grenzfrequenz angeordneten Harmonischen der Grundfrequenz bestimmt und neben der Grundfrequenz für eine Rekonstruktion des tieffrequenten Frequenzbereiches verwendet. Somit wird aus dem spektral ausgewerteten Abschnitt des Sprachsignals die maximale Information bezüglich des ungestörten Sprachsignals ausgenutzt. Die für die Rekonstruktion herangezogenen Frequenzen werden mit einer jeweiligen Spektralverteilung und einer vorgegebenen Amplitude zu einem synthetischen Spektrum zusammengesetzt, das den Frequenzbereich unterhalb der Grenzfrequenz im Sprachsignal entspricht. Aus diesem Frequenzabschnitt und dem Sprachsignal oberhalb der Grenzfrequenz wird dann das rekonstruierte Sprachsignal zusammengesetzt. Der tieffrequente Sprachanteil weist somit kein Rauschsignal mehr auf, da es ausschließlich aus Frequenzanteilen des Sprachsignals zusammengesetzt ist.

[0015] In einer weiteren Ausgestaltung der Erfindung kann der tieffrequente Sprachanteil auch direkt aus dem Sprachsignal ermittelt werden. Dazu wird ein aus mehreren Bandfiltern bestehendes Kammfilter auf der Basis der Grundfrequenz und der Frequenzen der unterhalb der Grenzfrequenz angeordneten Harmonischen eingerichtet, wobei die Frequenzpositionen der einzelnen Bandfilter den Grenzfrequenzen und der Harmonischen entsprechen. Mit Hilfe des Kammfilters wird dann das Sprachsignal im Bereich unterhalb der Grenzfrequenz gefiltert, wodurch die Signalanteile durchgelassen werden, die zum eigentlichen Sprachsignal gehören. Auch in dieser Weise ist eine Rekonstruktion eines weitgehend ungestörten Sprachsignals im tieffrequenten Bereich des Sprachsignals möglich.

[0016] Entscheidend für die Qualität der Rekonstruktion des tieffrequenten Sprachanteils ist die Genauigkeit der ermittelten Grundfrequenz des Sprachsignals. Da sich die Grundfrequenz während des Sprechens aufgrund der Satzmelodie laufend verändert, wird eine weitere Verbesserung des Verfahrens dadurch erreicht, daß zu Beginn eines Sprache enthaltenen Sprachabschnittes aus dem Sprachsignal die Grundfrequenz bestimmt wird und anschließend diese adaptiv nachgeführt wird. Somit wird im zeitlichen Verlauf des Sprachsignals jeweils die aktuelle Grundfrequenz bestimmt, so daß die Rekonstruktion des Sprachsignals möglichst genau an den Stimmverlauf angepaßt werden kann. Ein Ausführungsbeispiel einer solchen adaptiven Nachführung wird weiter unten im Detail erläutert.

[0017] In weiter bevorzugter Weise wird die Amplitu-

de des mindestens einen unterhalb der Grenzfrequenz erzeugten Frequenzsignals in Abhängigkeit von den Amplituden der oberhalb der Grenzfrequenz analysierten Frequenzsignale bestimmt. In weiter bevorzugter Weise können dabei typische Amplitudenverläufe von Sprachsignalen Anwendung finden, um nicht nur in den Frequenzanteilen, sondern auch in der Amplitudenverteilung der Frequenzanteile eine möglichst genaue Anpassung an ein natürliches Sprachsignal zu erreichen.

[0018] Weiter ist bevorzugt, daß die Grenzfrequenz in Abhängigkeit vom Geräuschpegel, also insbesondere von der Größe des Störsignals bestimmt wird. Somit ist es bei niedrigem Störsignalpegeln bspw. nur erforderlich, den Sprachsignalanteil unterhalb von 200 Hz zu rekonstruieren, während es bei hohen Störsignalpegeln notwendig ist, daß Sprachsignal im Frequenzbereich unterhalb von 500 Hz zu rekonstruieren. Bei einer Anwendung des Verfahrens in einem fahrenden Kraftfahrzeug kann die Grenzfrequenz auch in Abhängigkeit von der Fahrgeschwindigkeit bestimmt werden.

[0019] Weiterhin besteht eine Weiterbildung darin, daß das Sprachsignal vor einer Umwandlung einer Störsignalfreie unterzogen wird. Dabei können die herkömmlichen aus dem Stand der Technik bekannten Verfahren angewendet werden, um eine Vorbehandlung des Sprachsignals durchzuführen. Die Sprachanteile treten dann im Spektrum deutlicher hervor und können eindeutiger und somit genauer erkannt und rekonstruiert werden.

[0020] Eine Anwendung des zuvor beschriebenen Verfahrens besteht darin, in einem fahrenden Kraftfahrzeug aufgenommene Sprachsignale wiederzugeben, um dabei einen möglichst natürlichen Spracheindruck wiederzugeben.

[0021] Eine weitere Anwendung des erfindungsgemäßen Verfahrens besteht darin, ein mittels einer Telefonverbindung übertragenes Sprachsignal wiederzugeben. Das zugrunde liegende Problem besteht dabei darin, daß die Sprachsignale bei Telefonverbindungen im Frequenzbereich unterhalb von 350 Hz keine Informationen enthalten. Daher muß für eine naturgetreue Wiedergabe des Sprachsignals der tieffrequente Sprachanteil aus dem Frequenzbereich oberhalb von 350 Hz rekonstruiert werden. Dieses kann in besonders vorteilhafter Weise durch das erfindungsgemäße Verfahren durchgeführt werden.

[0022] Gemäß einer weiteren Lehre der vorliegenden Erfindung wird das oben dargestellte technische Problem auch durch eine Vorrichtung mit den Merkmalen des Anspruchs 12 gelöst, während in den Ansprüchen 13 bis 16 vorteilhafte Ausgestaltungen angegeben werden. Die Vorrichtung und das damit durchgeführte Verfahren werden im folgenden anhand von Ausführungsbeispielen näher erläutert, wobei auf die beigefügte Zeichnung Bezug genommen wird. In der Zeichnung zeigen

Fig. 1 eine spektrale Innengeräuschverteilung in ei-

- nem fahrenden Kraftfahrzeug für unterschiedliche Fahrgeschwindigkeiten,
- Fig. 2 ein Spektrogramm eines im tieffrequenten Bereich von einem Störsignal überlagerten Sprachsignals,
- Fig. 3 ein Spektrogramm des in Fig. 2 dargestellten Sprachsignals ohne Störsignal,
- Fig. 4 ein Spektrogramm des in Fig. 3 dargestellten Sprachsignals ohne Frequenzanteile unterhalb der Grenzfrequenz von ca. 400 Hz,
- Fig. 5 ein Spektrogramm der im Spektralbereich unterhalb der Grenzfrequenz von ca. 400 Hz rekonstruierten Sprachanteile,
- Fig. 6 das vollständige rekonstruierte Sprachsignal entsprechend dem in Fig. 3 dargestellten Sprachsignal ohne Störsignalanteil,
- Fig. 7 ein Blockschaltbild eines Ausführungsbeispiels einer erfindungsgemäßen Vorrichtung zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen,
- Fig. 8 eine Einrichtung zur adaptiven Nachführung der Grundfrequenz und
- Fig. 9 die spektrale Verteilung der Kennlinien der Bandfilter des Regelelementes zum Feststellen der frequenzabhängigen Leistungsverteilung im Mischspektrum in Bezug auf die feststehende Mischungsfrequenz von 2000 Hz.

[0023] In den Fig. 1 und 2 ist der Ausgangspunkt der vorliegenden Erfindung dargestellt.

[0024] Fig. 1 zeigt ein Frequenz-Amplituden-Diagramm des Innengeräuschpegels in einem fahrenden Kraftfahrzeug für unterschiedliche Geschwindigkeiten zwischen 60 Km/h und 160 Km/h. Bei dieser Darstellung fällt auf, daß insbesondere bei niedrigen Frequenzen unterhalb von ca. 500 Hz der Innengeräuschpegel im Vergleich zu den sonstigen Frequenzen des Innengeräuschsignals stark ansteigt. Da jedoch bei normaler Stimmlage die Grundfrequenz und die ersten Harmonischen zur Grundfrequenz im Frequenzbereich unter 1000 Hz und insbesondere unterhalb 500 Hz liegen, ist eine Bestimmung, also ein Herausfiltern des Sprachsignals aus dem Innenraumgeräuschsignal erheblich erschwert.

[0025] Fig. 2 zeigt ein Sprachsignal, das von einem Untergrundsignal überlagert worden ist, in einer Zeit-Frequenz-Darstellung als Spektrogramm. Dieses Spektrogramm wird bspw. durch eine Fouriertransformations (FFT) aus einem Mikrofonsignal erhalten. In Fig. 2 kennzeichnen unterschiedliche Grauwerte der Einzelseg-

mente des Spektrogramms unterschiedliche Intensitäten. Man erkennt einerseits deutlich die ansteigende Intensität (hellere Grauwerte) im Bereich kleiner Frequenzen zum Wert gleich Null hin und andererseits schmalbandige Frequenzanteile, die weitgehend parallel zueinander über kurze Zeitabschnitte verlaufen. Diese letztgenannten schmalbandigen Frequenzanteile stellen Harmonische der Grundfrequenz des entsprechenden Sprachsignals dar, die - wie im folgenden beschrieben - erfindungsgemäß ausgewertet werden.

[0026] Fig. 3 zeigt ein Spektrogramm des in Fig. 2 dargestellten Sprachsignals ohne das Untergrundgeräusch, so daß auch die tieffrequenten Sprachanteile als schmalbandige Frequenzanteile im Spektrogramm unterhalb von 500 Hz zu erkennen sind. Diese Sprachanteile gilt es zu rekonstruieren.

[0027] Fig. 4 zeigt weiterhin das zuvor dargestellte Sprachsignal, bei dem die Sprachanteile unterhalb einer Grenzfrequenz von ca. 400 Hz abgeschnitten sind. Ein derartiges Signal entspricht ungefähr dem Sprachsignal, wie es bei einer Telefonverbindung übertragen wird.

[0028] Fig. 5 zeigt ein Beispiel eines rekonstruierten Sprachsignals im Bereich unterhalb der Grenzfrequenz von ca. 400 Hz und Fig. 6 zeigt das zusammengesetzte rekonstruierte Sprachsignal aus dem in Fig. 5 dargestellten rekonstruierten Sprachanteil und dem in Fig. 4 dargestellten Frequenzanteil oberhalb der Grenzfrequenz des ursprünglichen Spektrums. Wie die rekonstruierten Sprachanteile erhalten werden, wird im folgenden anhand der Fig. 7 bis 9 im Detail beschrieben.

[0029] Fig. 7 zeigt in einem Blockschaltbild eine Vorrichtung zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen. Das Sprachsignal wird einem Mittel 4 zur Bestimmung von Frequenzanteilen ω_{fa1} , ω_{fa2} , ... von Maxima im Sprachsignal oberhalb einer vorgegebenen Grenzfrequenz ω_0 zugeleitet. Dazu wird das Sprachsignal zunächst durch ein Bandfilter 6 geleitet, so daß nur die Frequenzanteile zwischen der Grenzfrequenz ω_0 und einer weiteren Frequenz ω_1 herausgeschnitten und einer Weiterverarbeitung zugeleitet wird. ω_0 liegt dabei beispielsweise im Bereich von 200 bis 500 Hz, insbesondere bei 350 Hz, während die Frequenz ω_1 bspw. im Bereich von 800 Hz liegt. Der so ausgefilterte Frequenzabschnitt des Sprachsignals wird im Mischelement 8 gemischt, so daß die Summen- und Differenzfrequenzen der im herausgeschnittenen Abschnitt des Sprachsignals enthaltenen Frequenzanteile gebildet werden. Von Interesse sind dabei die Differenzfrequenzen, so daß das aus dem Mischelement 8 austretende Signal mittels eines Tiefpasses bearbeitet wird, so daß nur Frequenzanteile unterhalb einer einstellbaren Frequenz ω_2 durchgelassen werden. Somit läßt sich die kleinste Differenzfrequenz bestimmen, die dem Abstand zweier im Sprachsignal benachbart zueinander angeordneter Spektralanteile entspricht. Da es sich dabei um zwei Harmonische der Grundfrequenz handelt, stellt die Differenzfrequenz die Grundfrequenz ω_g dar.

Diese Grundfrequenz wird anschließend Mitteln 12 zur Rekonstruktion des Sprachsignals zugeleitet. Über einen weiteren Eingang der Mittel 12 wird das Sprachsignal über eine Verzögerungsstufe 14 und einen Tiefpaß 16 zugeführt. Somit liegt den Mitteln 12 sowohl der Wert der Grundfrequenz ω_g als auch ein vorgegebener Frequenzabschnitt des Sprachsignals für eine Rekonstruktion des die Sprache enthaltenden Signals zur Verfügung. Die Verzögerungsstufe 14 dient dabei einem Ausgleich der Zeitspanne Δt , die für die Bestimmung der Grundfrequenz ω_g benötigt wird und der Tiefpaß 16 dient einer sinnvollen Verringerung der Datenmenge, die den Mitteln 12 zur Rekonstruktion des Sprachsignals zugeleitet wird.

[0030] Die Mittel 12 zur Rekonstruktion des Sprachsignals unterhalb der Grenzfrequenz ω_0 weist schaltungstechnisch zwei Alternativen von Verfahrensweisen auf.

[0031] Als erste Alternative wird die Grundfrequenz ω_g herangezogen, um ein Signal im rekonstruierten Sprachsignal zu erzeugen, das dem Grundton der Sprache entspricht. Darüber hinaus können auch die Frequenzen der Harmonischen zur Grundfrequenz ω_g durch einfaches Multiplizieren mit den Zahlen $N = 2, 3, 4, \dots$ ermittelt werden, so daß für eine Rekonstruktion des Sprachanteils unterhalb der Grenzfrequenz ω_0 neben der Grundfrequenz ω_g auch die unterhalb der Grenzfrequenz ω_0 angeordneten Frequenzen $\omega_{h1}, \omega_{h2}, \dots$ der ersten, zweiten und weiteren Harmonischen verwendet werden. Ziel ist es dabei, sämtliche Harmonischen im zu rekonstruierenden Frequenzabschnitt des Sprachsignals zu erzeugen, also zu simulieren. Für eine spektrale Verteilung um jede dieser Frequenzen wird in Näherung eine Gauß'schen Verteilung oder eine andere mögliche spektrale Verteilung angenommen, die sich über eine Halbwertsbreite und eine Amplitude definieren läßt. Dadurch lassen sich die in Fig. 5 dargestellten spektralen Abschnitte im Spektrogramm erzeugen, die bei dem in Fig. 2 dargestellten verrauschten Signal nicht oder nur ansatzweise zu erkennen sind.

[0032] Als weitere Alternative für eine Rekonstruktion des tieffrequenten Sprachanteils besteht die Möglichkeit, daß die Mittel 12 einen Kammfilter aufweisen, der eine Mehrzahl von Bandfiltern aufweist, deren spektrale Durchlaßfunktionen durch die Grundfrequenz ω_g und die Frequenzen $\omega_{h1}, \omega_{h2}, \dots$ bestimmt werden. Die spektrale Durchlaßfunktion jedes Bandfilters wird zudem über eine vorgegebene Breite definiert, so daß entsprechende spektrale Abschnitte aus dem Sprachsignal im Bereich tiefer Frequenzen unterhalb der Grenzfrequenz ω_0 herausgefiltert werden. Da aus dem Spektrogramm nur die Anteile herausgefiltert werden, die das Sprachsignal enthalten, wird das Sprachsignal aus dem Spektrogramm rekonstruiert. Wird dabei zusätzlich eine Rauschunterdrückung durchgeführt, so werden aus den herausgefilterten Signalanteilen auch die Untergrundgeräusche herausgefiltert, so daß ein nahezu natürliches Sprachsignal erzeugt wird.

[0033] Wie weiterhin in Fig. 7 zu erkennen ist, wird das Sprachsignal über eine weitere Verzögerungsstufe 18 um eine Zeitdifferenz Δt verzögert, um eine Anpassung an die für Rekonstruktion des tieffrequenten Sprachanteils notwendige Zeitspanne zu ermöglichen. Nach Durchlaufen einen Hochpasses 20, in dem das Sprachsignal oberhalb der Grenzfrequenz ω_0 herausgefiltert wird, laufen sowohl dieses hochpaßgefilterte Signal als auch das rekonstruierte Sprachsignal für Frequenzen $\omega < \omega_0$ in dem Summenelement 22 zusammen, woraus das in Fig. 6 dargestellte rekonstruierte Spektrogramm erzeugt wird. Dieses Spektrogramm besteht also einerseits aus dem unterhalb der Grenzfrequenz ω_0 rekonstruierten Frequenzanteil sowie aus dem ursprünglichen Frequenzspektrum oberhalb der Grenzfrequenz ω_0 . Das so erzeugte Spektrogramm führt nach einer Umwandlung in ein Lautsprechersignal zu einer nahezu natürlich klingenden Sprachwiedergabe.

[0034] Wie bereits oben erläutert worden, bleibt im allgemeinen die Grundfrequenz ω_g in einem Sprachsignal aufgrund der Sprachmelodie nicht konstant. Daher ist es erforderlich, ständig die Grundfrequenz ω_g neu zu bestimmen. Dieses kann einerseits dadurch geschehen, daß ständig das zuvor beschriebenen Verfahren durchlaufen wird, das anhand der Elemente 4, 6, 8 und 10 zuvor beschrieben worden ist. Zum anderen kann jedoch eine genauere adaptive Nachführung der Grundfrequenz ω_g durchgeführt werden. Dieses ist mit einer Vorrichtung möglich, die in Fig. 8 dargestellt ist.

[0035] Die zu Beginn eines Sprachsignals zunächst bestimmte Grundfrequenz $\omega_{g,0}$ wird mit Hilfe eines Multiplikationselementes 24 auf den N-fachen Wert multipliziert. Somit wird die (N-1)te Harmonische zur Grundfrequenz berechnet. Die Frequenz dieser Harmonischen wird im folgenden als Regelharmonische bezeichnet und die zugehörige Frequenz mit ω_r bezeichnet.

[0036] Die Frequenz ω_r wird über einen Mehrtorschalter in einen Regelkreis eingebracht. In einer Initialisierungsphase zu Beginn eines Wortes wird der Ausgang des Multiplikationselementes 24 vom Mehrtorschalter 26 an das Mischelement 28 übergeben. Nach kurzer Zeit liegt - wie im folgenden beschrieben - ein Schätzwert ω_r neu vor und der Mehrtorschalter 26 wird so umgeschaltet, daß ω_r neu an das Mischelement 28 weitergegeben wird.

[0037] Ziel des Regelkreises besteht darin, die Differenz zwischen der (N-1)ten Harmonischen und einer festen Frequenz von bspw. $\omega_m = 2000$ Hz zu bestimmen. Im Idealfall ist ω_r exakt die Frequenz der (N-1)ten Harmonischen. Das Mischelement 28 bildet die Differenz zwischen ω_r und ω_m . Ein Sinusgenerator erzeugt ein sinusförmiges Signal mit der Frequenz, die durch sein Eingangssignal ω_d vorgegeben wird. Dieses wird einem Mischelement 32 zugeleitet, das das Sprachsignal und dieses sinusförmige Signal mischt. Nach erfolgter Mischung wird aus dem Mischelement 32 das gemischte Signal ausgegeben, das einem Regelelement 34 zum

Feststellen der frequenzabhängigen Leistungsverteilung im Mischsignal in Bezug auf die feststehende Frequenz ω_m zugeleitet wird.

[0038] Unter der Annahme, daß die dem Mischelement 28 zugeführte Frequenz ω_r der Regelharmonischen genau zu einer Harmonischen im aktuellen Sprachsignal paßt, entspricht die Summe aus der Differenzfrequenz ω_d , die durch die Differenz mit der feststehenden Mischfrequenz ω_m und ω_r erzeugt worden ist, und einem der Regelharmonischen entsprechenden Frequenzanteils des Sprachsignals genau der Mischfrequenz ω_m . Dieses spiegelt sich in einer Leistungsverteilung (P-Verteilung) im Leistungsspektrum wider. Die Leistungsverteilung wird bei der Mischfrequenz ω_m maximal sein.

[0039] Entspricht die Frequenz ω_r der Regelharmonischen jedoch nicht der aktuellen Frequenz der entsprechenden Harmonischen im Sprachsignal, so wird die Leistungsverteilung ihr Maximum nicht bei der Frequenz ω_m , sondern bei einer um einen Differenzwert $\Delta\omega$ verschobene Positionen annehmen. Somit läßt sich ein Korrekturwert zu $\Delta\omega$ bestimmen, der dem aktuellen Wert der Frequenz ω_r der Regelharmonischen hinzu addiert wird. Daraus entsteht der neue Wert der Frequenz $\omega_{r,neu}$, der über den Multiportschalter 26 der Regelschleife erneut zugeführt wird. Anschließend erfolgt erneut eine Mischung im Mischelement 28 mit nachfolgender Regelabfolge, wie sie zuvor beschrieben worden ist. Ändert sich somit im Laufe des Sprachsignals die Grundfrequenz und somit auch die Frequenz der entsprechenden Harmonischen im Sprachsignal, so wird dieses durch die Regelschleife ausgeglichen, so daß ständig ein aktueller, mit der Grundfrequenz ω_r weitgehend übereinstimmender Wert ω_r erzeugt.

[0040] Fig. 9 zeigt dazu die Kennlinien einer Mehrzahl von Bandfiltern, die für eine Bestimmung der Leistungsverteilung im Regelement 34 vorgesehen sind. Aus Fig. 9 ergibt sich eine Anzahl von 7 Bandfiltern, die um die feststehende Mischfrequenz $\omega_m = 2000$ Hz herum angeordnet sind. Fällt also beispielsweise die maximale Leistung in den Durchlaßbereich des mittleren Bandfilters, so wird der Korrekturwert $\Delta\omega=0$ gesetzt. Liegt dagegen das Maximum in einem der benachbart angeordneten Bandfilter, so wird ein entsprechender Korrekturwert $\Delta\omega \neq 0$ erzeugt, um bei weiter fortgeführter Regelung das Maximum der spektralen Leistungsverteilung in den Durchlaßbereich des mittleren Bandfilters zu verschieben.

[0041] Der Wert ω_r wird aus der Regelschleife über ein Multiplikationselement 38 abgezweigt und ausgegeben, in dem die aktuelle Frequenz ω_r mit dem Faktor $1/N$ beaufschlagt wird, um den Wert der Grundfrequenz $\omega_{g,adapt}$ zu erzeugen. Somit wird der Wert der Grundfrequenz ω_g ständig adaptiv nachgeführt, wodurch die Rekonstruktion des tieffrequenten Sprachanteils aus den mittelhohen Frequenzanteilen verbessert und näher an ein natürliches Sprachsignal herangeführt wird.

Patentansprüche

1. Verfahren zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen,
 - bei dem oberhalb einer Grenzfrequenz (ω_0) mindestens zwei benachbart angeordnete Frequenzanteile ($\omega_{fa1}, \omega_{fa2}, \dots$) mit erhöhter Amplitude im Sprachsignal bestimmt werden und
 - bei dem die Grundfrequenz (ω_g) des Sprachsignals als Frequenzdifferenz zwischen den mindestens zwei benachbarten Frequenzanteilen ($\omega_{fa1}, \omega_{fa2}, \dots$) bestimmt wird und
 - bei dem mit Hilfe der ermittelten Grundfrequenz (ω_g) und des Sprachsignals der tieffrequente Frequenzbereich unterhalb der Grenzfrequenz (ω_0) rekonstruiert wird.
2. Verfahren nach Anspruch 1, bei dem aus der Grundfrequenz (ω_g) die Frequenzen ($\omega_{h1}, \omega_{h2}, \dots$) der unterhalb der Grenzfrequenz (ω_0) angeordneten Harmonischen der Grundfrequenz (ω_g) bestimmt und neben der Grundfrequenz (ω_g) für das Rekonstruieren des tieffrequenten Frequenzbereiches verwendet werden.
3. Verfahren nach Anspruch 1, bei dem mit Hilfe eines mehrere Bandfilter aufweisenden Kammfilters auf der Basis der Grundfrequenz (ω_g) und der Frequenzen der unterhalb der Grenzfrequenz (ω_0) angeordneten Harmonischen die Frequenzpositionen der Bandfilter eingerichtet werden, mit deren Hilfe das Sprachsignals im Bereich unterhalb der Grenzfrequenz (ω_0) gefiltert wird.
4. Verfahren nach einem der Ansprüche 1 bis 3, bei dem zu Beginn eines Sprache enthaltenen Sprachabschnittes aus dem Sprachsignal die Grundfrequenz (ω_g) bestimmt wird und anschließend die Grundfrequenz (ω_g) adaptiv nachgeführt wird.
5. Verfahren nach Anspruch 4,
 - bei dem für eine adaptive Nachführung der Grundfrequenz (ω_g) aus dem aktuellen Wert der Grundfrequenz (ω_g) die Frequenz (ω_r) einer Regelharmonischen als N-te Harmonische berechnet wird,
 - bei dem die Differenz zwischen der Frequenz (ω_r) der Regelharmonischen und einer feststehenden Mischfrequenz (ω_m) gebildet wird,
 - bei dem ein sinusförmiges Signal ($\sin(\omega_d)$) mit der sich aus der Differenzbildung ergebenden Differenz- oder Summenfrequenz (ω_d) erzeugt wird,
 - bei dem das sinusförmige Signal ($\sin(\omega_d)$) mit dem Sprachsignal gemischt und ein Mischsignal erzeugt wird,

- bei dem im Mischsignal die frequenzabhängige Leistungsverteilung in Bezug auf die feststehende Mischungsfrequenz (ω_m) festgestellt wird,
 - bei dem aus der Leistungsverteilung ein Korrekturwert ($\Delta\omega$) für die Frequenz (ω_r) der Regelharmonischen berechnet wird,
 - bei dem die Frequenz (ω_r) der Regelharmonischen um den Korrekturwert ($\Delta\omega$) verändert und einer erneuten Mischung mit der feststehenden Mischungsfrequenz (ω_m) zugeleitet wird und
 - bei dem die Grundfrequenz (ω_g) ausgegeben wird, die dem entsprechenden Bruchteil $1/N$ der Frequenz (ω_r) entspricht.
6. Verfahren nach Anspruch 5, bei dem für eine Bestimmung die Leistungsverteilung das Mischsignal einer Mehrzahl von Bandfiltern (BF_n) zugeleitet wird, die nebeneinanderliegende Frequenzbereiche zentriert um die feste Mischungsfrequenz herum abdecken,
7. Verfahren nach einem der Ansprüche 1 bis 6, bei dem die Amplitude des mindestens einen unterhalb der Grenzfrequenz erzeugten Frequenzsignals in Abhängigkeit von den Amplituden der oberhalb der Grenzfrequenz analysierten Frequenzsignale bestimmt wird.
8. Verfahren nach einem der Ansprüche 1 bis 7, bei dem die Grenzfrequenz in Abhängigkeit vom Geräuschpegel bestimmt wird.
9. Verfahren nach einem der Ansprüche 1 bis 8, bei dem das Sprachsignal vor einer Umwandlung in ein Spektrogramm einer Störsignalfreie unterzogen wird.
10. Anwendung eines Verfahrens nach einem der Ansprüche 1 bis 9 für die Wiedergabe eines in einem fahrenden Kraftfahrzeug aufgenommenen Sprachsignals.
11. Anwendung eines Verfahrens nach einem der Ansprüche 1 bis 9 für die Wiedergabe eines Sprachsignals, das mittels einer Telefonverbindung übertragen wird.
12. Vorrichtung zur Rekonstruktion tieffrequenter Sprachanteile aus mittelhohen Frequenzanteilen, insbesondere zur Durchführung eines Verfahren nach einem der Ansprüche 1 bis 11,
- mit Mitteln (4) zur Bestimmung von Frequenzanteilen ($\omega_{fa1}, \omega_{fa2}, \dots$) von Maxima im Sprachsignal oberhalb einer vorgegebenen Grenzfrequenz (ω_0),
 - mit Mitteln (8) zum Mischen der Frequenzanteile ($\omega_{fa1}, \omega_{fa2}, \dots$) zur Bestimmung der Grundfrequenz (ω_g) des Sprachsignals als Differenzfrequenz zwischen jeweils zwei benachbarten Frequenzanteilen ($\omega_{fa1}, \omega_{fa2}, \dots$) und
 - mit Mitteln (12) zur Rekonstruktion des Sprachsignals unterhalb der Grenzfrequenz (ω_0) aus der ermittelten Grundfrequenz (ω_g) und dem Sprachsignal.
13. Vorrichtung nach Anspruch 12, dadurch gekennzeichnet, daß die Mittel (12) zur Rekonstruktion des Sprachsignals unterhalb der Grenzfrequenz (ω_0) das Spektrogramm aus der Grundfrequenz (ω_g) und den Frequenzen ($\omega_{h1}, \omega_{h2}, \dots$) der unterhalb der Grenzfrequenz (ω_0) angeordneten Harmonischen der Grundfrequenz (ω_g) mit einer vorgegebenen Spektralverteilung und einer vorgegebenen Amplitudenverteilung bestimmt.
14. Vorrichtung nach Anspruch 12, dadurch gekennzeichnet, daß die Mittel (12) einen Kammfilter mit einer Mehrzahl von Bandfiltern aufweisen, wobei die Frequenzen der Bandfilter auf der Basis der Grundfrequenz (ω_g) und ggf. ein oder mehrerer unterhalb der Grenzfrequenz (ω_0) angeordneter Harmonischer der Grundfrequenz (ω_g) einstellbar sind.
15. Vorrichtung nach einem der Ansprüche 12 bis 14, dadurch gekennzeichnet, daß für ein adaptives Nachführen der Grundfrequenz (ω_g) vorgesehen sind,
- ein Multiplikatorelement (24) zum Erzeugen der N-ten Harmonischen der Grundfrequenz als Frequenz (ω_r) einer Regelharmonischen,
 - ein Mischelement (28) zum Mischen der Frequenz (ω_r) der Regelharmonischen mit einer feststehenden Mischungsfrequenz (ω_m),
 - einem Sinusgenerator (30) zum Mischen der sich aus der Mischung ergebenden Differenz- oder Summenfrequenz (ω_d),
 - einem Mischelement (32) zum Mischen des sinusförmigen Signals ($\sin(\omega_d)$) mit dem Sprachsignal und zum Erzeugen eines Mischsignals erzeugt wird,
 - einem Regelelement (34) zum Feststellen der frequenzabhängigen Leistungsverteilung im Mischsignal in Bezug auf die feststehende Mischungsfrequenz (ω_m) und zum Berechnen eines Korrekturwert ($\Delta\omega$) für die Frequenz (ω_r) der Regelharmonischen aus der Leistungsverteilung,
 - einem Mischelement (36) zum Verändern der Frequenz (ω_r) der Regelharmonischen um den Korrekturwert ($\Delta\omega$) und
 - mit einem Multiplikatorelement (38) zum Berechnen des Bruchteils $1/N$ der Frequenz (ω_r)

als Grundfrequenz (ω_g).

16. Vorrichtung nach Anspruch 15, dadurch gekennzeichnet, daß das Regelement (34) eine Mehrzahl von Bandfiltern aufweist, die zentrisch zur Mischungsfrequenz (ω_m) nebeneinanderliegende Frequenzbereiche abdecken.

10

15

20

25

30

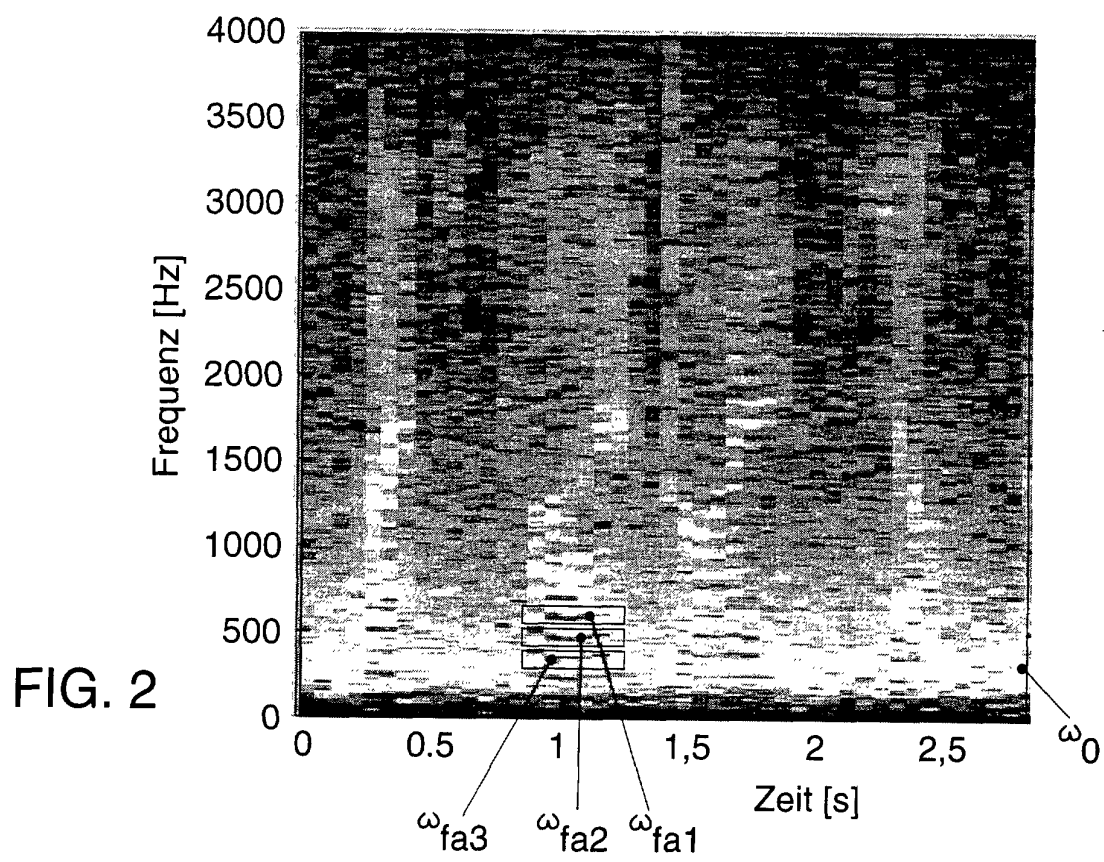
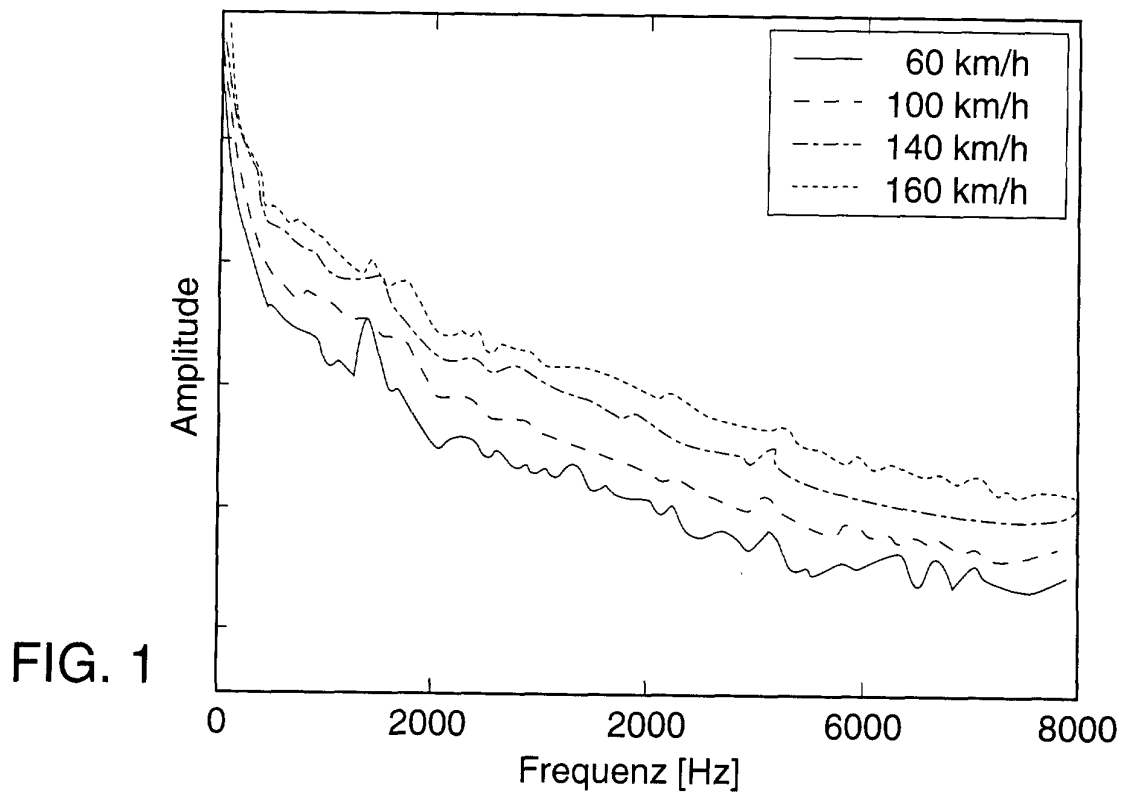
35

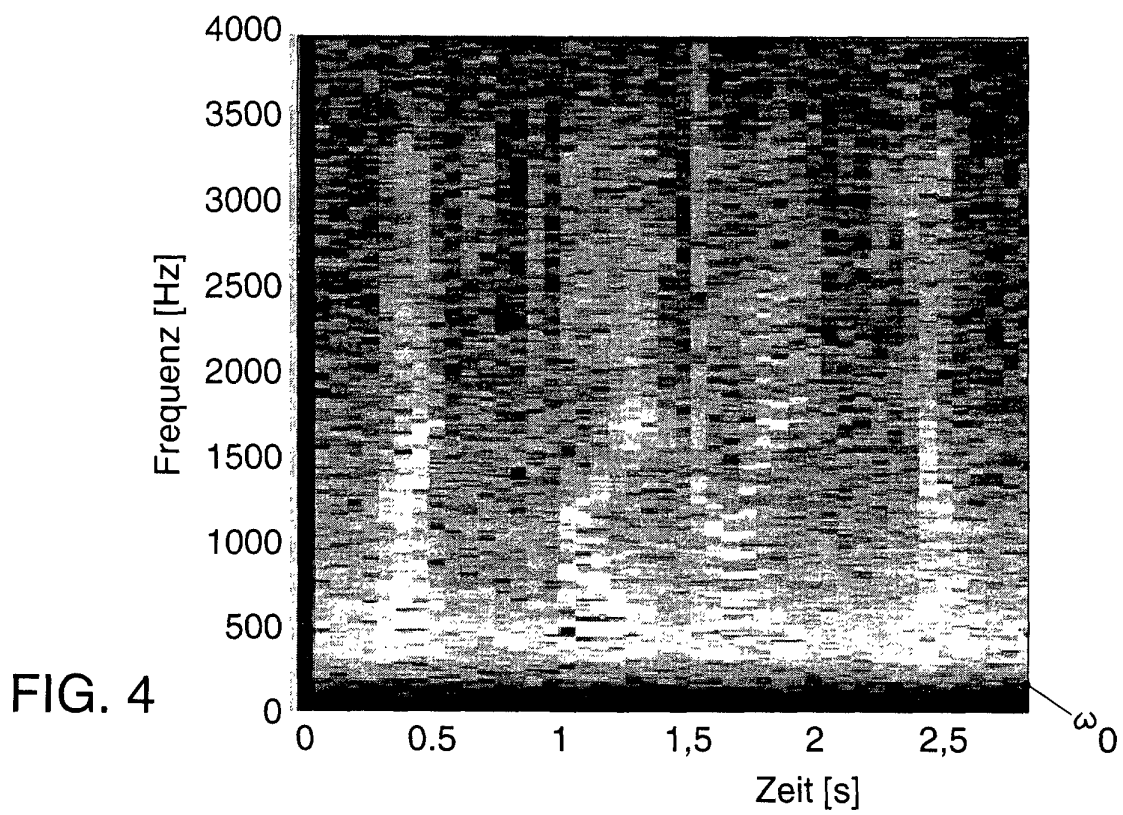
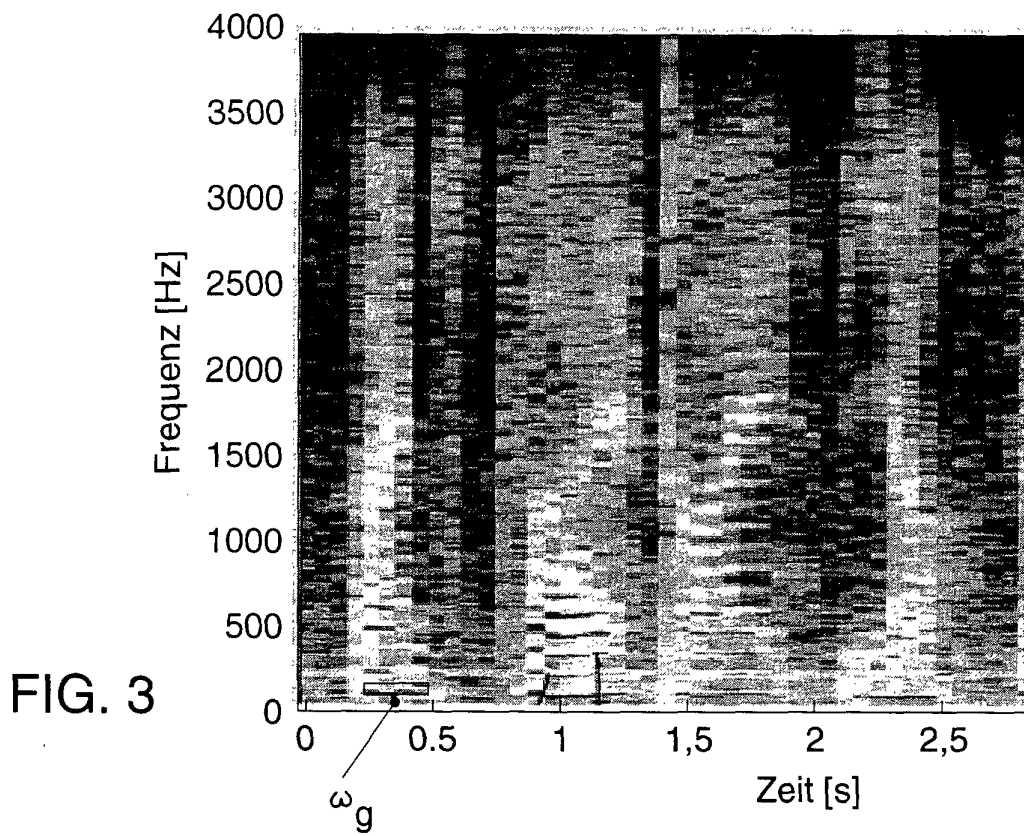
40

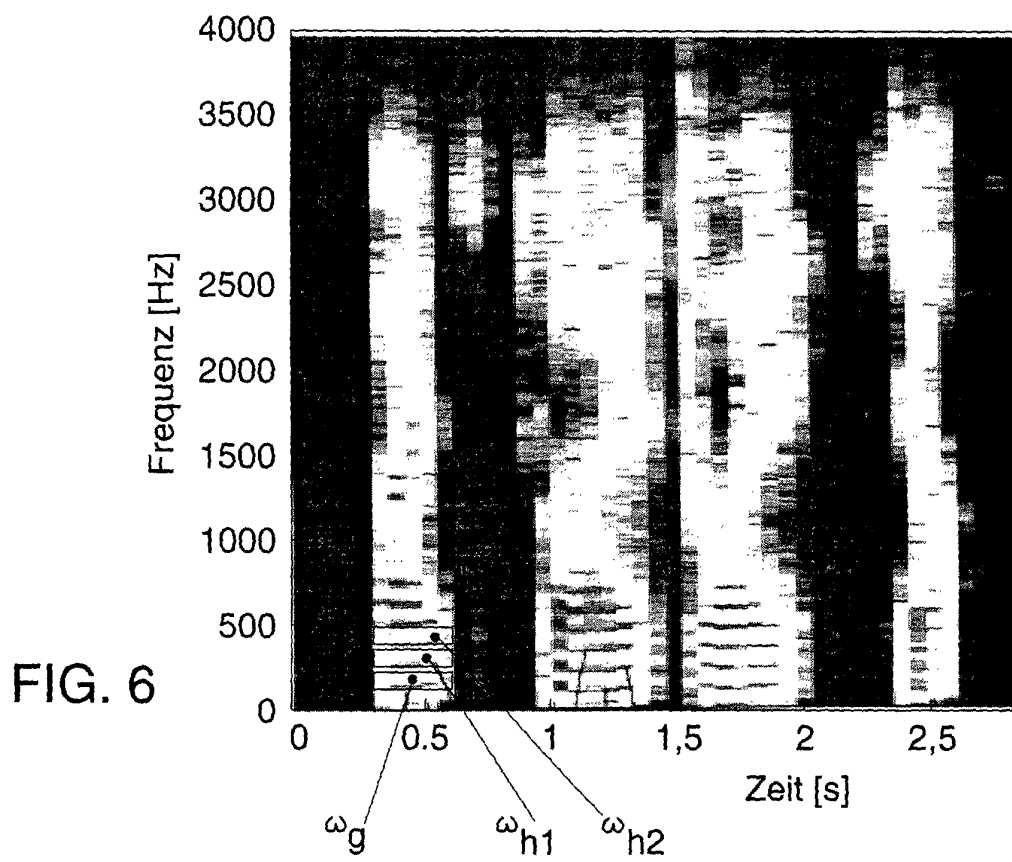
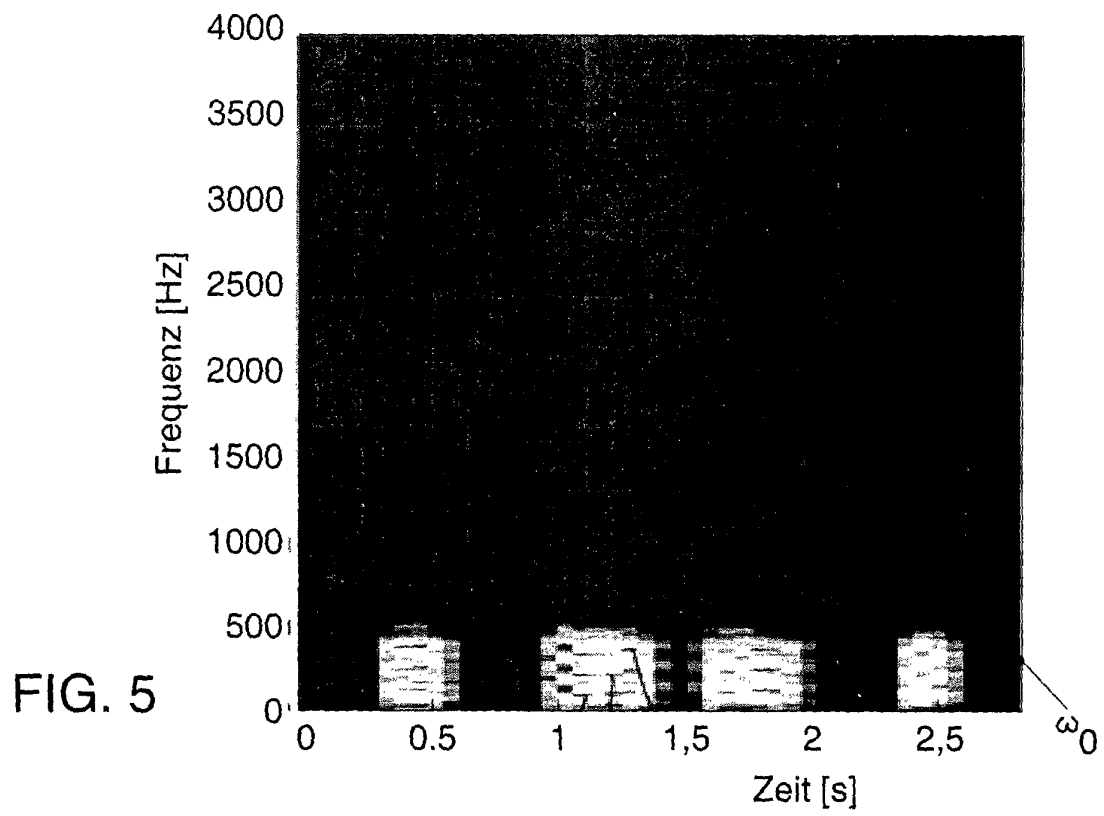
45

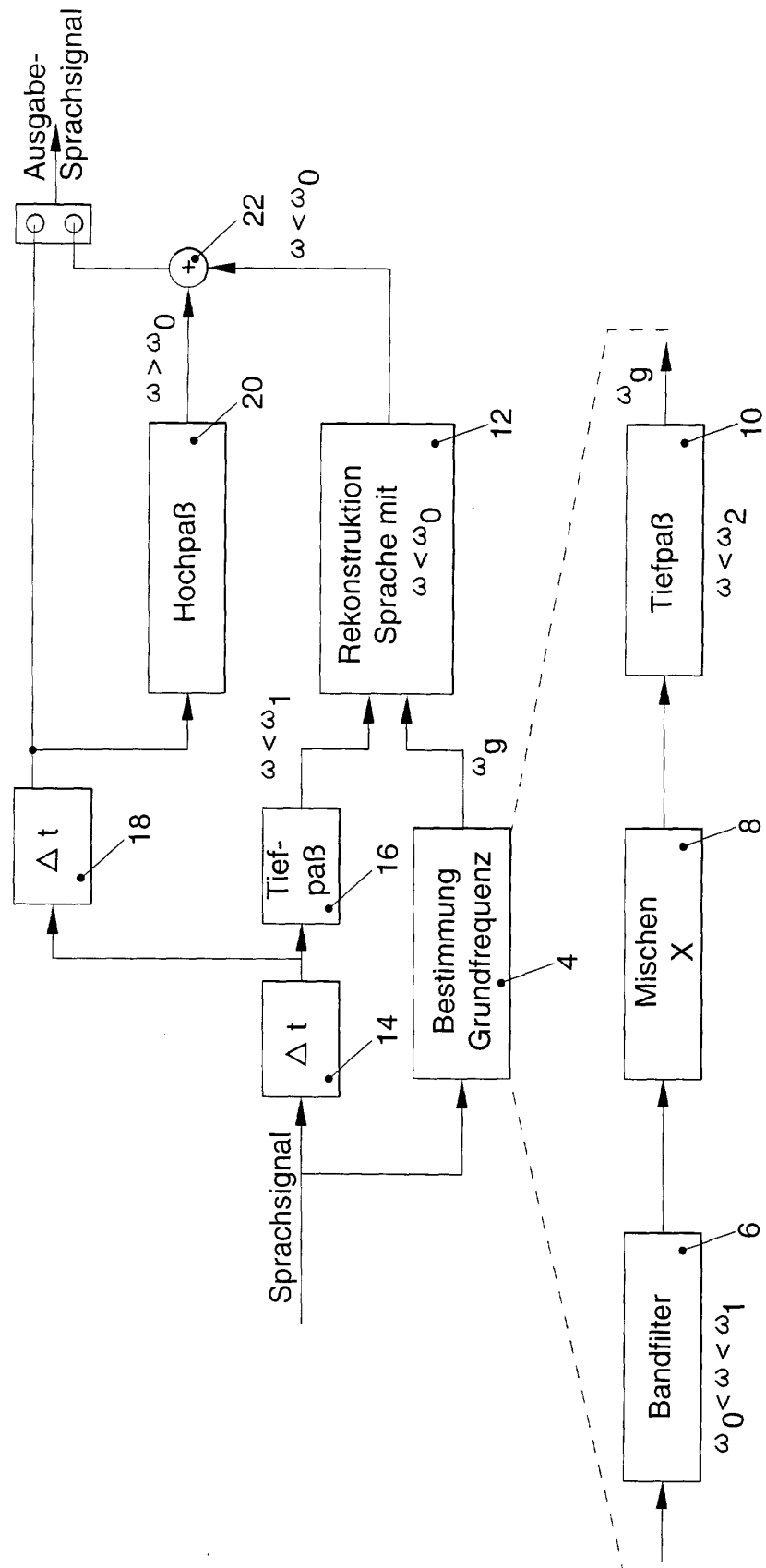
50

55









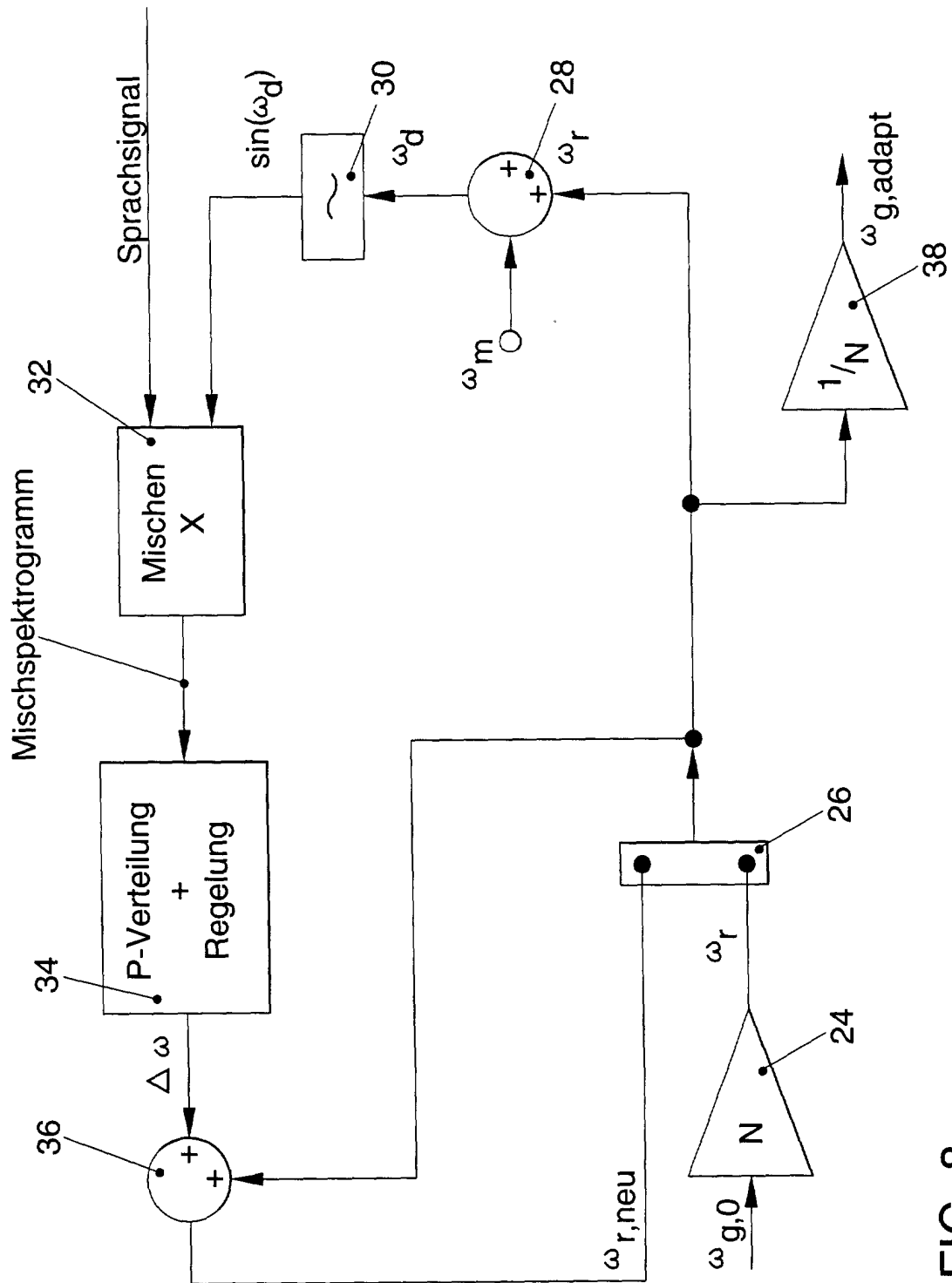


FIG. 8

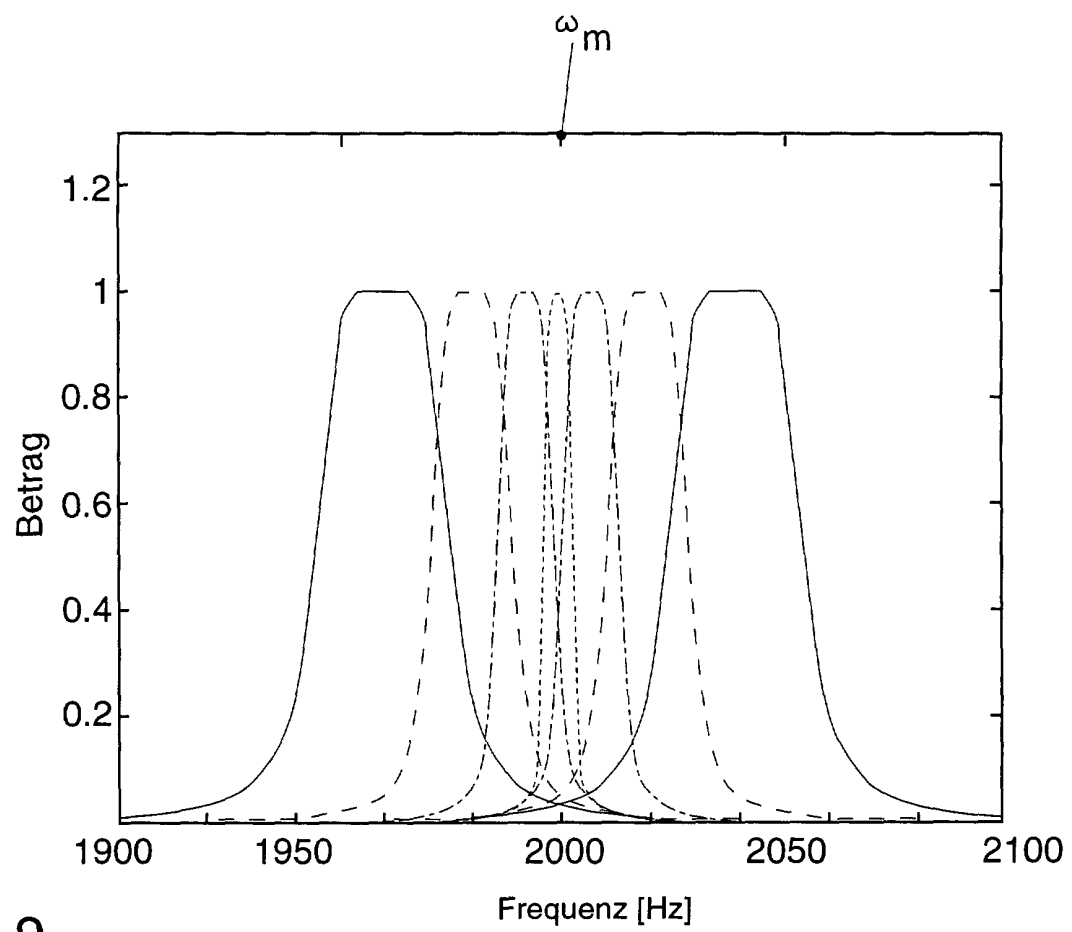


FIG. 9